May 20, 2005

LA-UR-04-4012

A Network Model of Early Events in Epidermal Growth Factor Receptor Signaling That Accounts for Combinatorial Complexity

Michael L. Blinov, James R. Faeder, Byron Goldstein, and William S. Hlavacek

Theoretical Biology and Biophysics Group, Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

Correspondence: Bill Hlavacek, Mail Stop K710, LANL, Los Alamos, NM 87545. Tel.: 505-665-1355; Fax: 505-665-3493; E-mail: wish@lanl.gov

Running title: A Network Model of EGFR Signaling

Keywords: computational systems biology, receptor tyrosine kinase, protein complexes, rule-based modeling, automatic network generation, BioNetGen

Abstract

We consider a model of early events in signaling by the epidermal growth factor (EGF) receptor (EGFR). The model includes EGF, EGFR, the adapter proteins Grb2 and Shc, and the guanine nucleotide exchange factor Sos, which is activated through EGF-induced formation of EGFR-Grb2-Sos and EGFR-Shc-Grb2-Sos assemblies at the plasma membrane. The protein interactions involved in signaling can potentially generate a diversity of protein complexes and phosphoforms; however, this diversity has been largely ignored in models of EGFR signaling. Here, we develop a model that accounts more fully for potential molecular diversity by specifying rules for protein interactions and then using these rules to generate a reaction network that includes all chemical species and reactions implied by the protein interactions. We obtain a model that predicts the dynamics of 356 molecular species, which are connected through 3,749 unidirectional reactions. This network model is compared with a previously developed model that includes only 18 chemical species but incorporates the same scope of protein interactions. The predictions of this model are reproduced by the network model, which also yields new predictions. For example, the network model predicts distinct temporal patterns of autophosphorylation for different tyrosine residues of EGFR. A comparison of the two models suggests experiments that could lead to mechanistic insights about competition among adapter proteins for EGFR binding sites and the role of EGFR monomers in signal transduction.

1 Introduction

Processes by which a cell senses and responds to its environment are often marked by combinatorial complexity (Hlavacek et al., 2003). Cellular signaling (Hunter, 2000) generally involves protein-protein interactions and enzymatic activities that imply a diversity of potential protein complexes and phosphoforms, which are difficult to simply enumerate let alone assay or understand. For example, the number of possible phosphoforms of a protein is 2^n, where n is the number of amino acid residues that are subject to phosphorylation and dephosphorylation by kinases and phosphatases, at least nine tyrosines in the case of epidermal growth factor (EGF) receptor (EGFR) (Jorissen et al., 2003). Additional molecular diversity can arise from the multivalent character of protein-protein interactions. A protein involved in signaling typically consists of multiple protein interaction domains (Pawson and Nash, 2003), such as the Src homology 2 (SH2) and 3 (SH3) domains of the Grb2 adapter protein. Each of these domains may interact with a partner that also contains multiple domains. As a result, proteins can combine in a variety of ways to form a spectrum of heterogeneous complexes. Proteomic studies confirm that diverse molecular species arise during signal transduction (Husi et al., 2000; Bunnell et al., 2002; Blagoev et al., 2003; 2004).

Given the protein-protein interactions and enzymatic activities involved in the cellular response to a signal, how do we catalog the potential molecular species implied by these interactions and activities? How do we predict which of the possible molecular species might actually arise during signaling? How do we determine the functional implications of these molecular species or the relative importance of processes that influence them? How can we best use large-scale proteomic measurements to obtain mechanistic insights? These questions are being asked in the emerging field of systems biology, and mathematical models have an important role to play in addressing such questions (Bhalla and Iyengar, 1999; Endy and Brent, 2001; Wiley et al., 2003; Hlavacek et al., 2003; Goldstein et al., 2004). A mathematical model requires an explicit statement of our understanding (or assumptions) about how a signal transduction system operates in a form that allows, through computational analysis, the behavior of the system to be predicted and compared with experimental observations. Here, we provide a demonstration of how a mathematical model,

incorporating detail at the level of protein sites and domains, can be used to study signal transduction with a comprehensive treatment of protein complexes and phosphoforms implied by protein interactions.

We develop and analyze a mathematical model for early events in signaling by EGFR, which is a well-studied cell-surface receptor involved in cell proliferation (Schlessinger, 2000; Jorissen et al., 2003). It has been the subject of numerous model-based studies (Wiley et al., 2003). Our model, which we will call the network model, provides a description of EGFR signaling that accounts for the spectrum of molecular species (356) and the reactions among these species (3,749) implied by specified interactions and activities of EGF, EGFR, the adapter proteins Grb2 and Shc, and the guanine nucleotide exchange factor Sos. These interactions and activities are the same as those considered in the seminal model of Kholodenko et al. (1999), which is based on assumptions (simultaneous phosphorylation and dephosphorylation of receptor tyrosines, inability of phosphorylated receptors in a dimer to dissociate, and competition among cytosolic proteins for receptor binding) that significantly limit, *a priori*, the number of molecular species that can arise during signaling. We will call the model of Kholodenko et al. (1999) the pathway-like model, because it represents the signaling system as a set of reaction sequences rather than as a highly branched reaction network.

The rest of this report is organized as follows. First, we describe how the network model is constructed based on the proteins, interactions, and model parameters considered in the work of Kholodenko et al. (1999). Notably, the network model involves no more parameters than the pathway-like model. We then compare the predictions of the two models with the experimental observations of Kholodenko et al. (1999). We find both models are equally consistent. We also present new predictions of the network model and testable predictions that distinguish the two models. A comparison of the models allows us to evaluate the simplifying assumptions of Kholodenko et al. (1999). These assumptions have not been tested so far, even though this model has served as the starting point for a number of modeling studies of EGFR signaling (Schoeberl et al., 2002; Gong and Zhao, 2003; Hatakeyama et al., 2003; Resat et al., 2003; Conzelmann et al.,

2004; Liu et al., 2005). We suggest experiments that could lead to insights into the mechanisms of signaling and determine which of the two models better represents signaling. Finally, we use the network model to predict the dynamics of the protein complexes and protein phosphorylation states that are generated during signaling. These predictions provide a picture of molecular diversity that is more detailed than could be currently obtained using the most sophisticated proteomic assays. For example, the model predicts which molecular species containing membrane-proximal Sos are prevalent at different time points. The model could also be used to predict how the population of these species depends on reaction dynamics and concentrations of components. As proteomic technologies mature, testing such predictions will become feasible.

2 The network model

2.1 Basis of the model

The network model (Figure 1) is based on the same proteins, enzymatic activities and protein-protein interactions considered in the model of Kholodenko et al. (1999). The focus of this model is the cascade of signaling events that lead to recruitment of cytosolic Sos to the inner cell membrane (Figures 1A and 1B), which can be described as follows. EGF binds to EGFR, which leads to the formation of signaling-competent receptor dimers (Garrett et al., 2002; Ogiso et al., 2002; Ferguson et al., 2003; Burgess et al., 2003). A receptor in a dimer then can be transphoshorylated by the tyrosine kinase domain of the neighboring receptor (Schlessinger, 2000; Jorissen et al., 2003). The cytosolic adapter proteins Grb2 and Shc are recruited to phosphorylated receptor tyrosines Y1068 and Y1148/73 (Batzer et al., 1994; Okabayashi et al., 1994). Shc is known to bind at two tyrosine residues, Y1148 via its SH2 domain and Y1173 via its phosphotyrosine binding (PTB) domain (Batzer et al., 1994). When Shc is bound to a receptor, it can be phosphorylated by EGFR (Pelicci et al., 1992). The phosphorylated form of Shc interacts with Grb2 (Lowenstein et al., 1992; Rozakis-Adcock et al., 1992; Sasaoka et al., 1994), which interacts constitutively with Sos (Egan et al., 1993; Rozakis-Adcock et al., 1993; Li et al., 1993). Translocation of Sos from the

cytosol to the membrane is required for activation of the GTPase Ras (Boguski and McCormick, 1993) and downstream signaling events. The model of Kholodenko et al. (1999) accounts for 25 species, including 18 involved in Sos recruitment.[1]

Some aspects of the model of Kholodenko et al. (1999) are controversial and might be subject to future investigation and refinement. For example, the model assumes that phosphorylation of Shc leads to a significant reduction in its affinity for EGFR, which is primarily responsible for the predicted damping of the initial response to EGF. Although recent molecular dynamics simulations support a lower affinity of phosphorylated Shc for EGFR (Suenaga et al., 2004), the implication that Shc recruitment and phosphorylation negatively regulates signaling is problematic in light of earlier experimental work on EGFR signaling (Sasaoka et al., 1994). In addition, pre-formed dimers of EGFR (Jorissen et al., 2003) and other complicating features of ligand-induced receptor dimerization that may influence signaling (Wofsy et al., 1992; Klein et al., 2004) are omitted in the model of Kholodenko et al. (1999) and its extensions. Our main focus here, however, is to evaluate the effects of simplifying assumptions made in developing the pathway-like model, and we therefore keep both the basic reaction processes and their accompanying rate constants in the network model so that we can make a controlled comparison of the two models.

We derive the network model without making the following assumptions, upon which the pathway-like model is based: 1) all receptor tyrosines of both receptors in a dimer are phosphorylated and dephosphorylated simultaneously, 2) dimeric receptors that are phosphorylated or associated with a cytosolic protein cannot dissociate, and 3) binding of cytosolic proteins to dimeric receptors is competitive, with only one cytosolic protein being allowed to associate with a dimer at a time. Without these assumptions, many more molecular species and reactions must be considered. Figure 1C illustrates one of the molecular species considered only in the network model, and Figure 1D illustrates the reactions in which this receptor dimer can participate. The network model is derived by applying a rule-based modeling approach (Hlavacek et al., 2003; Goldstein et al., 2004). Rules are specified based on the interactions and activities of protein domains. Each

[1]The model also includes the sequence of events leading to activation (phosphorylation) of phospholipase Cγ (PLCγ), which binds to phosphorylated Y992 and becomes transphosporylated by receptor tyrosine kinase.

rule defines a reaction class that is composed of multiple reactions that are parameterized by identical rate constants (Tables 1 and 2). This approach to parameterization is based on the assumption that protein domains and sites are modular. In other words, we assume that the activity of, say, the binding site on EGFR for Grb2 is independent of other sites on EGFR. A representative of each reaction class is illustrated in Figure 1A for reactions involving EGFR and in Figure 1B for cytosolic reactions. Each representative reaction corresponds with a "step" in the signaling cascade and with an individual reaction considered in the model of Kholodenko et al. (1999).

In Box 1 of Figure 1A, Step 1 is EGF binding to monomeric EGFR, and Step 2 is dimerization of EGF-bound receptors. In the network model, we allow modified receptors in dimers to dissociate, and as a result, there are multiple forms of monomeric EGFR available to bind EGF once signaling begins (but not before EGF stimulation). In contrast, in the pathway-like model, only receptors without modification participate in Steps 1 and 2, at all times. In fact, monomeric EGFR that is phosphorylated or associated with a cytosolic protein is not considered at all in the pathway-like model. To parameterize ligand-receptor and receptor-receptor interactions in the network model, we assume that the cytoplasmic state of a receptor does not affect ligand-receptor binding or ligand-stimulated receptor dimerization. Thus, we use the same rate constants estimated by Kholodenko et al. (1999) for all reactions involving the various possible states of receptors that can participate in Steps 1 and 2. Similar assumptions are made to parameterize the other steps in the signaling cascade (Tables 1 and 2). This approach to parameterization of the network model is the same approach taken to model early events in signaling by the immunoreceptor FcϵRI (Goldstein et al., 2002; Faeder et al., 2003; Goldstein et al., 2004).

The next steps in the intracellular signaling cascade are transphosphorylation of a receptor in a dimer by the tyrosine kinase domain of the neighboring EGFR (Step 3) and dephosphorylation (Step 4), which is carried out by phosphatases that are considered implicitly, as in the model of Kholodenko et al. (1999). In the pathway-like model, all receptor tyrosines of dimeric EGFR are lumped together, and Step 3 is considered to result in a single form of dimeric EGFR, designated RP, that encompasses, for example, the distinct receptor phosphoforms illustrated in Box 2 of

Figure 1A.

In contrast, we treat the phosphorylation sites of EGFR independently, and assume that there is no interaction among proteins binding to the different sites, which allows us to capture in the network model the full stoichiometric range of EGFR complexes without introducing any new rate parameters. We consider EGFR to contain one docking site for phospholipase Cγ (PLCγ), which is localized around Y992 and active when Y992 is phosphorylated, one docking site for Grb2, which is localized around Y1068 and active when Y1068 is phosphorylated, and one docking site for Shc, which is localized around Y1148 and Y1173 and active when tyrosines Y1148 and Y1173, which we lump together, are phosphorylated. These docking sites are assumed to be phosphorylated in separate reactions (via a non-processive mechanism), and association of PLCγ, Grb2, and Shc with EGFR is assumed to depend only on the phosphorylation states of Y992, Y1068, and Y1148/73, respectively. Association of a cytosolic protein (e.g., Grb2) with EGFR is assumed not to block association of a second cytosolic protein (e.g., Shc) with the same receptor or with an adjacent receptor in the same dimer.

Steps 5–8 and 25 in Figures 1A and 1B are related to the dynamics of PLCγ, which is considered in the model of Kholodenko et al. (1999) but not required for activation of Sos. Because we assume docking sites on EGFR are independent (see above), interaction of EGFR with PLCγ is unaffected by Grb2, Shc, or Sos. As a result, the models for activation (i.e., phosphorylation) of PLCγ and recruitment of Sos can be derived independently.[2]

Steps 9–24 in Figures 1A and 1B are related to the dynamics of Grb2, Shc, and Sos. Each of these steps is described briefly in Table 1. Translocation of Sos from the cytosol to the membrane is accomplished as follows. The cytosolic adapter proteins Grb2 and Shc are recruited to autophos-

[2]The model for activation of PLCγ, based on Boxes 1–3 in Figure 1A and Steps 8 and 25 in Figure 1B, accounts for 22 molecular species and consists of an equal number of ordinary differential equations (ODEs) describing the mass-action kinetics of these species. These species are enumerated as follows. Monomeric EGFR can be either free or bound to EGF. The PLCγ docking site on EGFR can be free, phosphorylated, bound to PLCγ, or bound to phosphorylated PLCγ. Thus, there are $2 \times 4 = 8$ possible forms of monomeric EGFR, four possible forms of symmetric receptor dimers (each receptor in a dimer is bound to EGF), and $\binom{4}{2} = 6$ possible forms of asymmetric receptor dimers. There are four additional possible species unassociated with EGFR, which are the same as those considered in the model of Kholodenko et al. (1999): 1) free EGF, 2) cytosolic resting PLCγ, 3) phosphorylated PLCγ in the cytosol (PLCγP), and 4) cytosolic PLCγ that has been inactivated and is no longer capable of interacting with EGFR (PLCγ-I). Parameters of this model are summarized in Table 2.

phorylated receptor tyrosines. Grb2 binds at Y1068 via its Src homology 2 (SH2) domain, and Shc binds at Y1148 via is SH2 domain and Y1173 via is phosphotyrosine binding (PTB) domain (Batzer et al., 1994; Okabayashi et al., 1994). Note that, just as we lump tyrosines Y1148 and Y1173 of EGFR together, we lump the SH2 and PTB domains of Shc together, treating Shc as if it has a single EGFR binding domain. When Shc is bound to a receptor, it can be phosphorylated by EGFR (Pelicci et al., 1992). The phosphorylated form of Shc interacts with the SH2 domain of Grb2 (Lowenstein et al., 1992; Rozakis-Adcock et al., 1992; Sasaoka et al., 1994), which interacts constitutively via its two SH3 domains with Sos (Egan et al., 1993; Rozakis-Adcock et al., 1993; Li et al., 1993). The reactions that take place in the cytosol (Steps 12, 16, and 21–23 of Figure 1B) are the same in the network and pathway-like models. The left-hand side of Box 4 in Figure 1A illustrates reactions involving EGFR that affect Shc-dependent recruitment of Sos to the membrane, and the right-hand side of Box 4 illustrates reactions involving EGFR that affect Shc-independent recruitment of Sos to the membrane.

2.2 Automatic generation of the reaction network

One can now enumerate the 356 molecular species considered in the network model for EGFR-mediated activation of Sos. The extracellular domain of EGFR can be free or bound to EGF. The Grb2 docking site of EGFR can be free, phosphorylated, bound to Grb2, or bound to Grb2 and Sos in complex. The Shc docking site of EGFR can be free, phosphorylated, bound to Shc, bound to phosphorylated Shc, bound to Shc and Grb2 in complex, or bound to Shc, Grb2, and Sos in complex. Thus, there are $2 \times 4 \times 6 = 48$ forms of monomeric EGFR, 24 forms of symmetric receptor dimers (each receptor in a dimer is bound to EGF), and $\binom{24}{2} = 276$ forms of asymmetric receptor dimers. There are eight additional species unassociated with EGFR: 1) free EGF, 2) cytosolic Grb2, 3) cytosolic Sos, 4) cytosolic Grb2 in complex with Sos (Grb2-Sos), 5) cytosolic Shc, 6) phosphorylated Shc in the cytosol (ShcP), 7) cytosolic Shc in complex with Grb2 (ShcP-Grb2), and 8) cytosolic Shc in complex with Grb2 and Sos (ShcP-Grb2-Sos).

Figure 1D illustrates 20 possible reactions a particular dimer can undergo: it can break up

into two receptor-monomers; unprotected phosphotyrosines can be dephosphorylated; phosphotyrosines can bind different cytosolic protein complexes; proteins can bind to proteins already associated with a receptor; some protein complexes may dissociate from receptor or from receptor-bound proteins. This example is typical. Table 1 gives the number of reactions of each reaction type included in the model. For example, a ligand can potentially bind to any receptor-monomer without a ligand, which gives rise to 48 unidirectional reactions; any two ligand-associated receptors can aggregate into a dimer, which gives rise to 600 potential dimerization reactions; etc. In the end, we consider 3,749 unidirectional reactions among the 356 species in the network model of Sos activation.

A computational model that describes interactions among these species typically includes an ordinary differential equation (ODE) for each chemical species. Each equation contains a term on the right-hand side for every reaction that influences the concentration of the species that corresponds to the ODE. To account for hundreds to thousands of species, one must abandon the approach of writing equations manually. An automatic procedure and software called BioNetGen (Blinov et al., 2004; Faeder et al., 2005a) was used to generate the chemical reaction network (and corresponding system of coupled ODEs) based on the reaction classes defined in Tables 1 and 2 and illustrated in Figure 1. The software and models used in our studies are available at our website (`http://cellsignaling.lanl.gov`). Given the relatively large size of the network model compared to most models of signal-transduction systems reported in the literature, it is reasonable to be concerned about the computational costs of a rule-based approach to modeling. The computational cost of integrating a system of 356 ODEs is manageable: computations using a standard method appropriate for stiff systems take at most a few minutes on a single processor. On the other hand, for larger systems, such as some of the simple extensions of the network model we have considered, which have tens of thousands of ODEs (unpublished material), computational efficiency is likely to be a major concern, and novel methods of simulation, in which rule evaluation is embedded in the method of simulation, may be needed (Lok and Brent, 2005; Faeder et al., 2005a).

Values of parameters in the network models for Sos and PLCγ activation (i.e., rate constants for each of the reaction classes and concentrations for EGF, EGFR, Grb2, Shc, Sos, and PLCγ) are summarized in Tables 1 and 2. These values correspond to parameter estimates of Kholodenko et al. (1999) with the following exceptions. Rate constants for reactions in which a cytosolic protein binds a docking site on EGFR have each been scaled by a factor of 0.5. This adjustment was made to account for the different stoichiometry of the two models: the pathway-like model accounts for only one docking site per receptor dimer, whereas the network model accounts for one docking site per receptor, or two docking sites per receptor dimer. Also, as a simplification, we have replaced Michaelis-Menten rate laws for phosphatase-catalyzed reactions with rate laws of the form kC, where k is a rate constant and C is the concentration of a phosphotyrosine-containing molecule. In each case, we have set the value of k equal to the value of the corresponding ratio V_M/K_M, which can be obtained from the estimates of Kholodenko et al. (1999). Note that V_M is the maximum velocity of a reaction, and K_M is the Michaelis-Menten constant. This simplification, which we have made in all the models that we consider, has an insignificant effect on results (results not shown). Schoeberl et al. (2002) also made this simplification.

3 Results

3.1 Similar predictions of the pathway-like and network models

The pathway-like model was parameterized so that its predictions would match certain dynamic responses to EGF that were assayed by Kholodenko et al. (1999). In Figures 2A–2E, the experimental results of Kholodenko et al. (1999) are compared with predictions of the network model and two forms of the pathway-like model: one with PLCγ included, as in the original model, and a reduced form with PLCγ omitted. We consider these two forms of the pathway-like model to illustrate that competitive binding of PLCγ, which is omitted in the network model, has a minimal effect on activation of Sos. As can be seen, for the quantities measured by Kholodenko et al. (1999), the network and two pathway-like models make similar predictions, especially at steady

state. The different models are more-or-less equally consistent with the measured time courses of direct Grb2 association with EGFR (Figure 2A), PLCγ phosphorylation (Figure 2B), indirect Grb2 association with EGFR via Shc (Figure 2C), Shc phosphorylation (Figure 2D), and receptor phosphorylation (Figure 2E). The differences at early times in Figures 2A–2D are explained by the assumption that cytosolic proteins compete for receptor binding in both forms of the pathway-like model, and the differences of Figure 2E are explained by the assumption of the pathway-like model that all receptor phosphotyrosines are protected from phosphatase activity if a single cytosolic protein is associated with a receptor dimer. After the transient, the models agree, largely because Shc and PLCγ are found mostly in their phosphorylated cytosolic forms at steady state (not shown), and these forms have low affinity for docking sites on EGFR (Kholodenko et al., 1999; Tables 1 and 2). Also, the dynamics of cytosolic reactions are identical in the network and pathway-like models. Because binding of cytosolic proteins to EGFR is non-competitive in the network model, the predictions of the network model shown in Figure 2 are independent of whether PLCγ is considered or omitted.

3.2 New predictions of the network model

The network model, because it incorporates more molecular details than the pathway-like model, can predict the results of experiments that are beyond the scope of the pathway-like model (Figure 3). Figure 3A shows that individual tyrosines of EGFR are predicted by the network model to display distinct dynamics during a response to EGF, even though the rise and fall of total receptor phosphorylation is predicted by the network and pathway-like models alike. Comparison of Figure 3A with Figure 2E reveals that the pattern of total receptor phosphorylation can be attributed to phosphorylation of Y1148/73, binding of Shc at pY1148/73, where it protects pY1148/73 from phosphatase activity, phosphorylation of Shc, and return of phosphorylated Shc to the cytosol, where it concentrates in its phosphorylated form because of the low affinity of phosphorylated Shc for EFGR and the slow rate of cytosolic Shc dephosphorylation in the model (Kholodenko et al., 1999; Table 1).

The remaining panels of Figure 3 illustrate that the network model predicts greater molecular diversity than the pathway-like model. Figure 3B shows that a substantial fraction of receptor dimers are predicted by the network model to associate transiently with two molecules of phosphorylated Shc during the response to EGFR. At the peak of the time course of Figure 3B, 26% of total phosphorylated Shc is found in complexes containing two molecules of EGFR and two molecules of Shc. Figure 3C shows that a substantial fraction of monomeric receptors are also predicted by the network model to be found in association with Sos. The slightly delayed association of Sos with monomeric receptors compared to association of Sos with dimeric receptors reflects the time required for receptors in dimers to become phosphorylated and dissociate from each other. As can be seen, the fraction of membrane-associated Sos in complexes with monomeric receptors is small at short times, but becomes more significant as the steady state is approached. At steady state, 36% of total Sos is associated with monomeric receptors. The assumptions of the pathway-like model preclude the formation of these complexes and many others, as indicated in Figure 3D. The plots of Figure 3D represent the numbers of molecular species containing 98% of receptors as a function of time after EGF stimulation according to the network and pathway-like models. As can be seen, at the peak of the time course predicted by the network model, 98% of receptors are distributed among more than 100 molecular species. The diversity of molecular species containing 98% of receptors falls as the steady state is approached to about one third of its peak transient value. However, at all times, the distribution of receptors among molecular species predicted by the network model is greater than that predicted by the pathway-like model. The molecular species (and reactions) that are prevalent at any given time are determined by the dynamics of the signaling cascade and the connectivity of the reaction network (Faeder et al., 2005b).

3.3 Distinguishing predictions of the two models

The two models make distinct predictions, which could help us determine if assumptions restricting the range of species are valid (Figure 4). Figure 4A illustrates the predicted effects of two mutations: 1) knocking out the EGFR tyrosines that bind Shc, and 2) knocking out Shc tyrosines that

bind Grb2. The first mutation blocks Shc recruitment to receptor. The second mutation blocks Grb2 binding to Shc, but allows Shc to bind EGFR. Both mutations eliminate Shc-mediated recruitment of Sos and thus might be expected to downregulate Sos activation. Figure 4A demonstrates that both mutations do indeed lower Sos recruitment compared to the wild type (WT) during transient activation, but, counter to intuition, produce a much higher level of steady-state Sos activation. For the first mutation, both the network and pathway-like models predict the same steady-state levels, because in the absence of Shc binding, Grb2 recruits Sos at the same rate in both models via direct binding of Grb2 to EGFR. For the second mutation, the predictions of the two models differ because unphosphorylated Shc can bind EGFR, and this binding competes with Grb2-EGFR binding in the pathway-like model but not in the network model. This competition results in partial inhibition of Sos activation in the pathway-like model, although the level of steady-state activation is still higher in the mutant than in WT. In contrast, the level of steady-state activation predicted by the network model is the same for both mutations. If the levels of steady-state Sos activation for the two mutants could be measured experimentally, the results presented above suggest that independence of the Grb2 and Shc binding sites in EGFR could be established or refuted.

The two models make different predictions about EGFR dimerization during signaling, as shown in Figure 4B. The pathway-like model predicts a slower approach to steady state and a larger number of receptors in dimers at steady state than the network model. The pathway-like model also predicts a transient overshoot, whereas the network model does not. These predictions of the pathway-like model arise, because in this model signaling affects receptor dimerization. Recall that phosphorylated receptors in a dimer are not allowed to dissociate. In contrast, in the network model, receptor dimerization is independent of the signaling events under consideration. Thus, monitoring receptor dimerization as parameters of signaling are varied (e.g., protein expression levels of Grb2 or Shc) could help determine if signaling affects dimerization. The predicted effect of varying Grb2 concentration is shown in Figure 4B. The number of receptor dimers depends on the total cytosolic Grb2 concentration in the pathway-like model but not in the network model.

The two models also make different predictions about the outcome of the following hypothetical experiment, which depends on whether or not receptor monomers are involved in signaling. Let us consider a panel of bivalent antibodies that crosslink EGFR with different kinetics. Figure 4C illustrates the predicted effects at steady state of crosslinking receptors with an antibody as a function of dimer dissociation rate. The network model predicts that Shc recruitment is maximal at an optimal dimer lifetime, whereas the pathway-like model predicts a low level of Shc recruitment that is relatively insensitive to the lifetime of dimers. In the pathway-like model only dimers can recruit Shc, whereas in the network model phosphorylated receptor-monomers can be associated with Shc. Model parameters are such that unphosphorylated Shc has high affinity for EGFR and remains associated with receptor for a long time, which explains the build-up of the receptor-associated pool of Shc proteins as dimer lifetime decreases. A shorter lifetime allows more receptors to traffic through the dimeric state, become phosphorylated and recruit Shc. This effect is one of serial engagement (Goldstein et al., 2004). As the lifetime becomes shorter still, Shc recruitment falls, because dimers break apart before receptors can be phosphorylated, an effect known as kinetic proofreading (Hlavacek et al., 2002; Goldstein et al., 2004).

3.4 Molecular diversity during signaling

In Figure 3D, we pointed out that only a relatively small fraction of the possible molecular species is populated at steady state, but many more species are populated during the transient peak. This behavior is described in greater detail in Figure 5, which helps to clarify the differences and similarities between the network and pathway-like models. The distribution of Sos in different species is shown at short times in panel A and at long times in panel B. Figure 5B demonstrates that at steady state only two chemical species contain 72% of all Sos recruited through ShcP. These species are a dimer and a monomer that each contain one phosphorylated tyrosine. This kind of species is accounted for in the model of Kholodenko et al. (1999). In contrast with the narrow distribution at steady state, the transient distribution is much broader. The two most prevalent steady state species contain only 47% of Sos recruited through ShcP (Figure 5A), with Sos distributed

among a variety of other receptor species: 26 species account for 95% of Sos recruited through ShcP. The next two most prevalent receptor species during the transient contain 18% of Sos at 10 sec. These species contain receptor dimers with each receptor bound to an adapter; this kind of species is unaccounted for in the pathway-like model. Similar distributions are observed for Sos bound to receptor via Grb2 recruited at Y1068 (Figures 5C and 5D).

These observations provide an explanation for the match in quantitative predictions between the two models at steady state. According to the network model, under steady state conditions, only one phosphotyrosine in a dimer is engaged in protein-binding activity, as assumed in the pathway-like model of Kholodenko et al. (1999). Complexes in which more than one protein is bound to a receptor at the same time are rare at steady state, but common during the transient, which explains the differences between the two models at short times.

3.5 Distinct reaction sequences that lead to Sos recruitment

Another important question is what reaction paths are involved in recruitment of Sos to the membrane. To answer this question we use the method of path analysis, described by Faeder et al. (2005b). This method determines the prevalence of sequences of reactions by which a molecular component is transformed from one state into another one. Here, we focus on pathways that transform an unmodified receptor monomer into a receptor associated with Sos under steady state conditions.

There are nine distinct reaction paths for Sos recruitment in the model of Kholodenko et al. (1999). Figure 6A illustrates the top three, which account for 70% of Sos recruited at steady state. In contrast, there is a variety of paths in the network model, with the 50 most prevalent paths accounting only for 45% of Sos recruitment. All of the paths in the network model can be grouped into 9 classes of reaction paths, each of which corresponds to an individual path in the model of Kholodenko et al. (1999). Figure 6B illustrates the most prevalent paths within one such class, in which recruitment of Grb2 and Sos is sequential.

According to both models, at steady state, about a third of Sos is activated via sequential

recruitment of Grb2 and Sos. Parameters of the models are such that the amount of free Grb2-Sos complex in the cytosol is less than the amount of free cytosolic Grb2, which explains why simultaneous recruitment of Grb2 and Sos in complex, the second most prevalent route to Sos activation, is less common than sequential recruitment of Grb2 and Sos to Y1068. These two classes of paths are responsible for about 50% of the Sos recruited at steady state. The third most prevalent class of paths is binding of the complex ShcP-Grb2-Sos to Y1148/73. Although the amount of ShcP-Grb2-Sos complexes is high at steady state, the prevalence of this pathway is supressed by the low affinity of the complex for EGFR.

As illustrated in Figure 6B, within each class of Sos activation pathways there is molecular diversity. The network model allows one to identify the individual chemical species involved in these reaction paths. As can be seen, sequences that recruit Sos may involve receptor monomers.

Gong and Zhao (2003) have also analyzed the relative contributions of the Shc-dependent and Shc-independent pathways for Sos recruitment and subsequent Ras activation using the model of Schoeberl et al. (2002), and they find that over the full time course of receptor stimulation there is a significant preference for the Shc-dependent pathway. We find for both the pathway-like and the network models that the two pathways make nearly equal contributions to Sos activation at steady state. The differences between our results and those of Gong and Zhao (2003) might arise from the differences between the model of Kholodenko et al. (1999) and that of Schoeberl et al. (2002), such as the assumption that GAP stabilizes the formation of dimers in the latter model. In any event, the results of Sec. 3.3 demonstrate that pathway analysis is not always a reliable indicator of the effect of blocking the activity of a signaling protein, which in the case of Shc results in a large increase in the level of steady-state Sos activation. Thus, when Shc is present it may carry a major portion of the Sos activation flux, but its removal, at least in the case of the model of Kholodenko et al. (1999), permits the Shc-independent pathway to carry a substantially larger activation flux.

4 Discussion

The problem of combinatorial complexity has been noted by a number of researchers (Husi et al., 2000; Arkin, 2001; Endy and Brent, 2002; Pawson and Nash, 2003). Nevertheless, very few modelers have attempted to account for the full potential molecular diversity of signaling systems, although models incorporating detail at the submolecular level have been developed (Wofsy et al., 1997; Morton-Firth and Bray, 1998; Levchenko et al., 1999; Goldstein et al., 2001; Faeder et al., 2002; Shimizu et al., 2003; Li et al., 2004, Haugh et al., 2004), including some that account for hundreds of species. Two problems hinder modeling of systems that involve multistate molecules and multicomponent complexes: tracking the combinatorially large number of possible molecular species, and defining their properties. We address these problems with a rule-based modeling approach, implemented in software called BioNetGen (Blinov et al., 2004; Faeder et al., 2005a). Computer-aided evaluation of rules, specified on the basis of knowledge of protein interactions, generates a reaction network, with the individual chemical species and reactions in the network defined automatically.

Here, we have presented a specific application of the rule-based approach for developing predictive models of signaling systems without ignoring combinatorial complexity. As demonstrated, a rule-based model can be derived from the same type of experimental data that has been used to construct models via the conventional manual approach, which is capable of accounting for only a fraction of the species that are potentially generated in a signaling system. In contrast, in the rule-based approach, the full spectrum of protein complexes and phosphoforms implied by specified molecular interactions and activities of the domains of the signaling molecules is considered. Thus, for example, covalent modification of each of the tyrosines of EGFR included in the network model presented here can be monitored independently. If each domain of a protein has modular binding or enzymatic activity, a protein consisting of several domains and/or sites of covalent modifications may undergo many possible reactions. Multiple branch points arise at each signaling step and lead to different possible signaling scenarios. All of these branch points can be considered in a rule-based model, because the number of domain interactions of interest is often small, which

allows corresponding rules (but not equations) to be specified manually.

Interactions among proteins in the network model considered in this paper are inherited from the interactions considered in the pathway-like model of Kholodenko et al. (1999). The pathway-like model represents early events in EGFR signaling as a series of consecutive signaling events and closely mirrors the hierarchical ordering of steps in the signaling cascade initiated by EGF stimulation. Variables in the model can be associated with experimental observables and reactions in the model correspond to steps in the signaling cascade. The model includes only 25 chemical species and 47 reactions, because it is based on a number of assumptions that limit molecular diversity *a priori*. These assumptions include simultaneous phosphorylation of all tyrosine residues in a dimer (i.e. lumping all the tyrosines together), no dissociation of phosphorylated receptor dimers, and competitive binding of adapter proteins to a receptor dimer. These assumptions may or may not be appropriate under all experimental conditions. We would like to reduce such limiting assumptions to a minimum even though we recognize that not all possible species are important under all conditions. The reason is that we cannot identify the critical species through intuition, as the importance of species depends on network dynamics and connectivity. Input to a model in the form of known molecular interactions and activities seems preferable to input in the form of an intuitively selected set of chemical species and reactions.

The network model, without adjustment of parameter values, reproduces the predictions of the pathway-like model that were compared with experimental data by Kholodenko et al. (1999) (Figure 2). However, other predictions of the two models differ, and the network model yields new predictions that are unavailable from the pathway-like model. For example, in the network model, individual tyrosines of EGFR are considered. Thus, this model can be used to predict the dynamics of phosphorylation for each tyrosine, and we find that different tyrosines have different temporal patterns of phosphorylation (Figure 3A). The results of Figure 4 suggest three types of experiments that could be performed to determine which model better represents early events in EGFR signaling and obtain mechanistic insights. For example, the pathway-like model predicts that mutations of EGFR and Shc tyrosines should have different effects, whereas the network

model predicts that these mutations should have identical effects (Figure 4A). Interestingly, both models make the surprising prediction that elimination of Shc-mediated recruitment of Sos will have a positive effect on overall Sos recruitment at steady state. This result is obtained because phosphorylated Shc in the models, consistent with molecular dynamics simulations (Suenaga et al., 2004), binds EGFR with low affinity and sequesters Sos in the cytosol, inhibiting Sos activation. This finding appears to contradict earlier experimental results (Sasaoka et al., 1994) and probably merits further experimental investigation.

Kinetic rate constants in the network model are the same as those in the pathway-like model with the exception of a scale factor, which was introduced for stoichiometric reasons (cf. Tables 1 and 2 and Kholodenko et al., 1999). For example, the rate constant for Grb2 binding to phosphorylated receptor dimers in the pathway-like model is twice the value of the rate constant for binding of the SH2 domain of Grb2 to receptor phosphotyrosine Y1068 in the network model. To avoid introducing additional rate parameters when constructing the network model, we made the assumption that proteins bind to the phosphorylation sites of EGFR independently, which creates large classes of reactions that can be characterized by the same rate constant. This lumping of rate constants in the network model is motivated by the modularity of protein domains (see Hlavacek et al., 2003), and it can be tested and refined as more experimental data becomes available about the effects of protein-protein interactions. There is, however, already some evidence that a large number of proteins can bind independently to the relatively small cytoplasmic portion of a membrane-bound receptor, such as EGFR (Jiang and Sorkin, 2002) or CD19 (Brooks et al., 2004).

The simple structure of the pathway-like model, on the other hand, comes at the cost of introducing complex and often hidden correlations among model parameters and variables. For example, the assumption that only unmodified dimers of EGFR can dissociate leads to a correlation between adapter protein expression levels and ligand-receptor binding (Figure 4B), and masks the effects of serial engagement and kinetic proofreading on EGFR-Shc association predicted by the network model (Figure 4C).

The pathway-like and network models provide different microscopic pictures of EGFR sig-

naling. As might be expected, the network model predicts greater molecular diversity (Figures 3B-3D, 5, and 6). For example, multiple adapter proteins are predicted to associate with the same receptor dimer, which is not allowed in the pathway-like model but consistent with the results of Jiang and Sorkin (2002). Also, receptor monomers are predicted to contribute to Sos recruitment, which again is not allowed in the pathway-like model. Spatial spread of receptor phosphorylation in response to localized EGF stimulation (Verveer et al., 2000; Reynolds et al., 2003) might be explained by the involvement of receptor monomers in signaling. Also consistent with predictions of the network model are proteomic assays indicating that EGFR is a member of a large number of heterogeneous protein complexes that arise during signaling (Blagoev et al., 2003; 2004). At this time, predictions about molecular diversity are difficult to test, but such predictions will become more important with advances in methods for monitoring protein modifications and interactions (Aebersold and Mann, 2003; Mann and Jensen, 2003; Meyer and Teruel, 2003).

Acknowledgements

This work was supported by DOE through contract W-7405-ENG-36 and by NIH through grants GM35556, AI28433 and RR18754. We thank Ed Stites for stimulating discussions.

References

Aebersold, R., Mann, M., 2003. Mass spectrometry-based proteomics. Nature 422, 198–207.

Arkin, A.P., 2001. Synthetic cell biology. Curr. Opinion Biotechnol. 12, 638–644.

Batzer, A.G., Rotin, D., Ureña, J.M., Skolnik, E.Y., Schlessinger, J., 1994. Hierarchy of binding sites for Grb2 and Shc on the epidermal growth factor receptor. Mol. Cell Biol. 14, 5192–5201.

Bhalla, U.S., Iyengar, R., 1999. Emergent properties of networks of biological signaling pathways. Science 283, 381–387.

Blagoev, B., Kratchmarova, I., Ong, S-E., Nielsen, M., Foster, L.J., Mann, M., 2003. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. Nat. Biotechnol. 21, 315–318.

Blagoev, B., Ong, S-E., Kratchmarova, I., Mann, M., 2004. Temporal analysis of phosphotyrosine-dependent signaling networks by quantitative proteomics. Nat. Biotechnol. 22, 1139–1145.

Blinov, M.L., Faeder, J.R., Goldstein, B., Hlavacek, W.S., 2004. BioNetGen: software for rule-based modeling of signal transduction based on the interactions of molecular domains. Bioinformatics 17, 3289–3291

Boguski, M.S., McCormick, F., 1993. Proteins regulating Ras and its relatives. Nature 366, 643–654.

Brooks, S. R., Kirkham, P. M., Freeberg, L., Carter, R. H., 2004. Binding of cytoplasmic proteins to the CD19 intracellular domain is high affinity, competitive, and multimeric. J. Immunol. 172, 7556-7564.

Bunnell, S.C., Hong, D.I., Kardon, J.R., Yamazaki, T., McGlade, C.J., Barr, V.A., Samelson, L.E., 2002. T cell receptor ligation induces the formation of dynamically regulated signaling assemblies. J. Cell Biol. 158, 1263–1275.

Burgess, A.W., Cho, H.S., Eigenbrot, C., Ferguson, K.M., Garrett, T.P.J., Leahy, D.J., Lemmon, M.A., Sliwkowski, M.X., Ward, C.W., Yokoyama, S., 2003. An open-and-shut case? Recent insights into the activation of EGF/ErbB receptors. Mol. Cell 12, 541–552

Chook, Y.M., Gish, G.D., Kay, C.M., Pai, E.F., Pawson, T., 1996. The Grb2-mSos1 complex binds phosphopeptides with higher affinity than Grb2. J. Biol. Chem. 271, 30472–30478.

Conzelmann, H., Saez-Rodriguez, J., Sauter, T., Bullinger, E., Allgower, F., Gilles, E.D., 2004. Reduction of mathematical models of signal transduction networks: simulation-based approach applied to EGF receptor signalling. Syst. Biol. 1, 159–169

Egan, S.E., Giddings, B.W., Brooks, M.W., Buday, L., Sizeland, A.M., Weinberg, R.A., 1993. Association of Sos Ras exchange protein with Grb2 is implicated in tyrosine kinase signal transduction and transformation. Nature 363, 45–51.

Emlet, D.R., Moscatello, D.K., Ludlow, L.B., Wong, A.J., 1997. Subsets of epidermal growth factor receptors during activation and endocytosis. J. Biol. Chem. 272, 4079–4086

Endy, D., Brent, R., 2001. Modeling cellular behavior. Nature 409, 391–395.

Faeder, J.R., Hlavacek, W.S., Reischl, I., Blinov, M.L., Metzger, H., Redondo, A., Wofsy, C., Goldstein, B., 2003. Investigation of early events in FcϵRI-mediated signaling using a detailed mathematical model. J. Immunol. 170, 3769–3781.

Faeder, J.R., Blinov, M.L., Goldstein, B., Hlavacek, W.S., 2005a. Rule-based modeling of biochemical networks. Complexity 10, 22–41.

Faeder, J.R., Blinov, M.L., Hlavacek, W.S., Goldstein, B., 2005b. Combinatorial complexity and dynamical restriction of network flows in signal transduction. Syst. Biol. 2, 5–15.

Ferguson, K.M., Berger, M.B., Mendrola, J.M., Cho, H-S., Leahy, D.J., Lemmon, M.A., 2003. EGF activates its receptor by removing interactions that autoinhibit ectodomain dimerization. Mol. Cell 11, 507–517.

Garrett, T.P.J., McKern, N.M., Lou, M., Elleman, T.C., Adams, T.E., Lovrecz, G.O., Zhu, H-J., Walker, F., Frenkel, M.J., Hoyne, P.A., Jorissen, R.N., Nice, E.C., Burgess, A.W., Ward, C.W., 2002. Crystal structure of a truncated epidermal growth factor receptor extracellular domain bound to transforming growth factor α. Cell 110, 763–773.

Gladhaug, I.P., Christoffersen, T., 1988. Rapid constitutive internalization and externalization of epidermal growth factor receptors in isolated rat hepatocytes. Monensin inhibits receptor externalization and reduces the capacity for continued endocytosis of epidermal growth factor. J. Biol. Chem. 263, 12199-12203.

Goldstein, B., Faeder, J.R., Hlavacek, W.S., 2004. Mathematical and computational models of immune-receptor signalling. Nat. Rev. Immunol. 4, 445–456

Goldstein, B., Faeder, J.R., Hlavacek, W.S., Blinov, M.L., Redondo, A., Wofsy, C., 2002. Modeling the early signaling events mediated by FcϵRI. Mol. Immunol. 38, 1213–1219.

Gong, Y.C., Zhao, X., 2003. Shc-dependent pathway is redundant but dominant in MAPK cascade activation by EGF receptors: A modeling inference. FEBS Lett. 554, 467–472

Hatakeyama, M., Kimura, S., Naka, T., Kawasaki, T., Yumoto, N., Ichikawa, M., Kim, J.H., Saito, K., Saeki, M., Shirouzu, M., et. al., 2003. A computational model on the modulation of mitogen-activated protein kinase (MAPK) and Akt pathways in heregulin-induced ErbB signalling. Biochem. J. 373, 451-463

Haugh, J.M., Lauffenburger, D.A., 1997. Physical modulation of intracellular signaling processes by locational regulation. Biophys. J. 72, 2014–2031.

Haugh, J.M., Schneider, I.C., Lewis, J.M. On the cross-regulation of protein tyrosine phosphatases and receptor tyrosine kinases in intracellular signaling J. Theor. Biol. 230, 119–132.

Hlavacek, W.S., Faeder, J.R., Blinov, M.L., Perelson, A.S., Goldstein, B., 2003. The complexity of complexes in signal transduction. Biotech. Bioeng. 84, 783–794

Hlavacek, W.S., Redondo, A., Wofsy, C., Goldstein, B., 2002. Kinetic proofreading in receptor-mediated transduction of cellular signals: Receptor aggregation, partially activated receptors, and cytosolic messengers. Bull. Math. Biol. 64, 887–911

Hunter, T., 1995. Protein kinases and phosphatases: the Yin and Yang of protein phosphorylation and signaling. Cell 80, 225–236.

Hunter, T., 2000. Signaling—2000 and beyond. Cell 100, 113–127.

Husi, H., Ward, M.A., Choudhary, J.S., Blackstock, W.P., Grant, S.G.N., 2000. Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. Nat. Neurosci. 3, 661–669.

Jiang, X., Sorkin, A., 2002. Coordinated traffic of Grb2 and Ras during epidermal growth factor receptor endocytosis visualized in living cells. Mol. Biol. Cell 13, 1522–1535.

Jorissen, R.N., Walker, F., Pouliot, N., Garrett, T.P.J., Ward, C.W., Burgess, A.W., 2003. Epidermal growth factor receptor: mechanisms of activation and signalling. Exp. Cell Res. 284, 31–53.

Kholodenko, B.N., Demin, O.V., Moehren, G., Hoek, J.B., 1999. Quantification of short term signaling by the epidermal growth factor receptor. J. Biol. Chem. 274, 30169–30181.

Klein, P., Mattoon, D., Lemmon, M.A., Schlessinger, J., 2004. A structure-based model for ligand binding and dimerization of EGF receptors. Proc. Natl. Ac. of Sci. 101, .929–993

Levchenko, A., Bruck, J,, Sternberg, P.W., 2000. Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties. Proc. Natl. Acad. Sci. USA 97, 5818–5823.

Li, N., Batzer, A.G., Daly, R.J., Yajnik, V., Skolnik, E.Y., Chardin, P., Bar-Sagi, D., Margolis, B., Schlessinger, J., 1993. Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinase to Ras signalling. Nature 363, 85–88.

Li, Q.J., Dinner, A.R., Qi, S.Y., Irvine, D.J., Huppa, J.B., Davis, M.M., Chakraborty, A.K. CD4 enhances T cell sensitivity to antigen by coordinating Lck accumulation at the immunological synapse Nat. Immunol. 5, 791–799.

Liu, G., Swihart, M.T., Neelamegham S., 2005. Sensitivity, principal component and flux analysis applied to signal transduction: the case of epidermal growth factor mediated signaling. Bioinformatics 21, 1194–1202.

Lok, L., Brent, R., 2005. Automatic generation of cellular reaction networks with Moleculizer 1.0. Nat. Biotechnol. 23, 131–136.

Lowenstein, E.J., Daly, R.J., Batzer, A.G., Li, W., Margolis, B., Lammers, R., Ullrich, A., Skolnik, E.Y., Bar-Sagi, D., Schlessinger, J., 1992. The SH2 and SH3 domain containing protein GRB2 links receptor tyrosine kinases to ras signaling. Cell 70, 431–442.

Mann, M., Jensen, O.N., 2003. Proteomic analysis of post-translational modifications. Nat. Biotechnol. 21, 255–261.

Meyer, T., Teruel, M.N., 2003. Fluorescence imaging of signaling networks. Trends Cell Biol. 13, 101–106.

Morton-Firth, C.J., Bray, D., 1998. Predicting temporal fluctuations in an intracellular signalling pathway. J. Theor. Biol. 192, 117–128.

Ogiso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J-H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., Yokoyama, S., 2002. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. Cell 110, 775–787.

Okabayashi, Y., Kido, Y., Okutani, T., Sugimoto, Y., Sakaguchi, K., Kasuga, M., 1994. Tyrosines 1148 and 1173 of activated human epidermal growth factor receptors are binding sites of Shc in intact cells. J. Biol. Chem. 269, 18674–18678.

O'Rourke, L., Ladbury, J.E., 2003. Specificity is complex and time consuming: mutual exclusivity in tyrosine kinase-mediated signaling. Acc. Chem. Res. 36, 410–416.

Pawson, T., Nash, P., 2003. Assembly of cell regulatory systems through protein interaction domains. Science 300, 445–452.

Pelicci, G., Lanfrancone, L., Grignani, F., McGlade, J., Cavallo, F., Forni, G., Nicoletti, I., Grignani, F., Pawson, T., Pelicci, P.G., 1992. A novel transforming protein (SHC) with an SH2 domain is implicated in mitogenic signal transduction. Cell 70, 93–104.

Resat, H., Ewald, J.A., Dixon, D.A., Wiley, H.S., 2003. An integrated model of epidermal growth factor receptor trafficking and signal transduction. Biophys. J. 85, 730-743.

Reynolds, A.R., Tischer, C., Verveer, P.J., Rocks, O., Bastiaens, P.I.H., 2003. EGFR activation coupled to inhibition of tyrosine phosphatases causes lateral signal propagation. Nat. Cell Biol. 5, 447–453

Rozakis-Adcock, M., Fernley, R., Wade, J., Pawson, T., Bowtell, D., 1993. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. Nature 363, 83–85.

Rozakis-Adcock, M., McGlade, J., Mbamalu, G., Pelicci, G., Daly, R.J., Li, W., Batzer, A.G., Thomas, S., Brugge, J., Pelicci, P.G., Schlessinger, J., Pawson, T., 1992. Association of the Shc and Grb2/Sem5 SH2-containing proteins is implicated in activation of the Ras pathway by tyrosine kinases. Nature 360, 689–692.

Sasaoka, T., Langlois, W.J., Leitner, J.W., Draznin, B., Olefsky, J.M., 1994. The signaling pathway coupling epidermal growth factor receptors to activation of p21^{ras}. J. Biol. Chem. 269,

32621–32625.

Schlessinger, J., 2000. Cell signaling by receptor tyrosine kinases. Cell 103, 211–225.

Schoeberl, B., Eichler-Jonsson, C., Gilles, E.D., Müller G., 2002. Computational modeling of the dynamics of the MAP kinase cascade activated by surface and internalized EGF receptors. Nat. Biotechnol. 20, 370–375.

Shimizu, T., Aksenov, S.V., Bray, D., 2003. A spatially extended stochastic model of the bacterial chemotaxis signalling pathway. J. Mol. Biol. 329, 291–309.

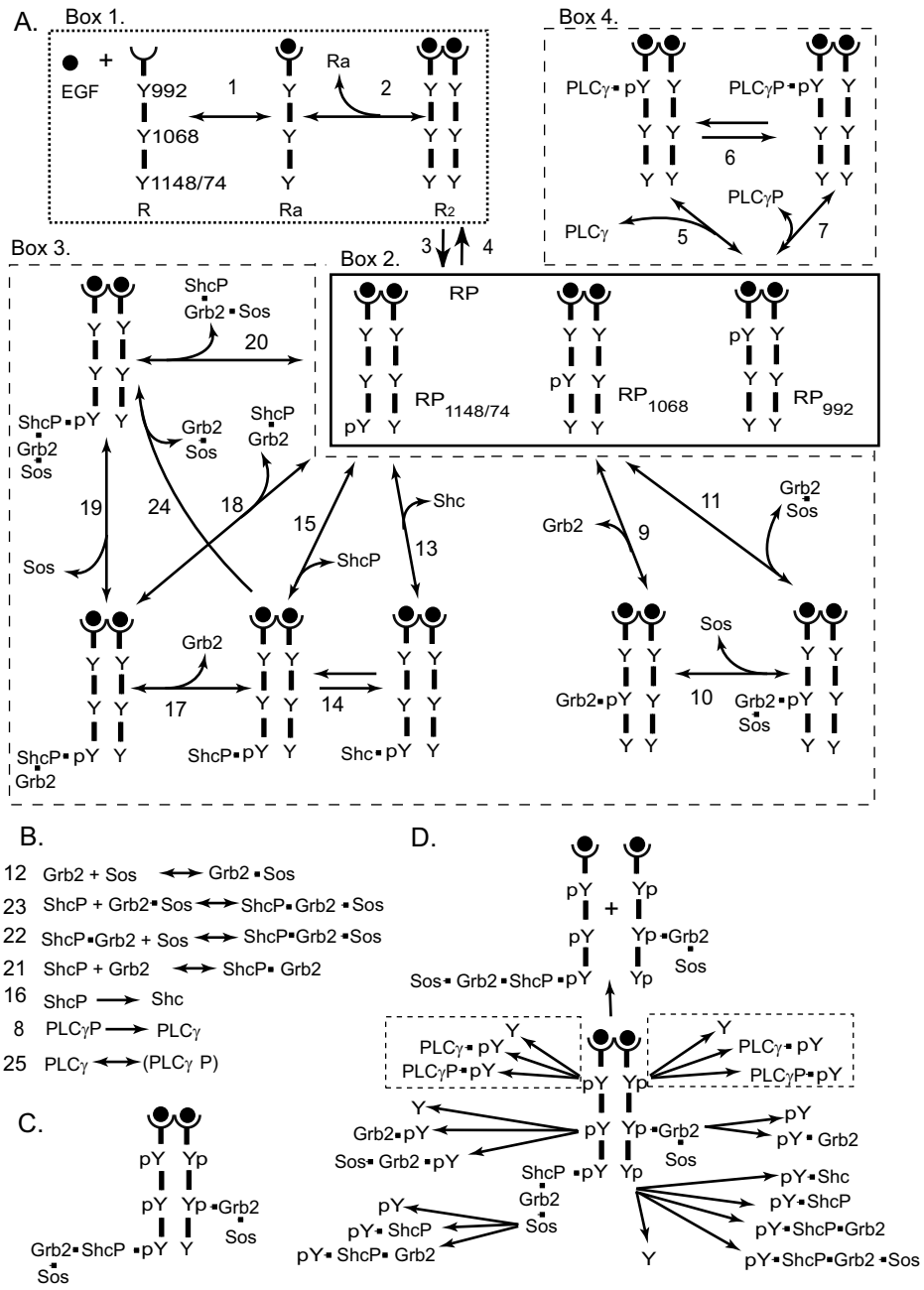
Suenaga, A., Kiyatkin, A.B., Hatakeyama, M., Futatsugi, N., Okimoto, N., Hirano, Y., Narumi, T., Kawai, A., Susukita, R., Koishi, T., et. al., 2004. Tyr-317 phosphorylation increases Shc structural rigidity and reduces coupling of domain motions remote from the phosphorylation site as revealed by molecular dynamics simulations. J. Biol. Chem. 279, 4657–4646

Verveer, P.J., Wouters, F.S., Reynolds, A.R., Bastiaens, P.I.H., 2000. Quantitative imaging of lateral ErbB1 receptor signal propagation in the plasma membrane. Science 290, 1567–1570

Wiley, H.S., Shvartsman, S.Y., Lauffenburger, D.A., 2003. Computational modeling of the EGF-receptor system: a paradigm for systems biology. Trends Cell Biol. 13, 43–50.

Wofsy, C., Goldstein, B., Lund, K., Wiley, H.S., 1992. Implications of Epidermal Growth Factor (EGF) induced EGF receptor aggregation. Biophys. J. 63, 98–110

Figure 1. Kinetic scheme for the early events in EGFR signaling. (A) Reactions involving receptors. Only EGFR tyrosine resides that are required for interactions with PLCγ, Grb2 and Shc are included. EGF-EGFR binding and receptor dimerization reactions are illustrated inside Box 1. Although only two reactions are illustrated, EGF may bind a receptor in any cytoplasmic state, and any two EGF-bound receptors can aggregate or dissociate at any time. After receptor aggregation different receptor phosphoforms may be formed as a result of receptor tyrosine kinase activity. Three of the possible phosphoforms are illustrated and named inside Box 2. Cytoplasmic binding reactions related to Sos recruitment and PLCγ activation are shown inside Boxes 3 and 4, respectively. Each reaction is an example of many possible reactions. (B) Cytosolic reactions not involving receptors. In panels A and B, numbers next to each reaction refer to signaling steps and reaction classes described in Table 1. (C) Example of a species omitted from consideration in the model of Kholodenko et al. (1999). (D) Illustration of 20 possible reactions the dimeric species of panel C may undergo in the network model. In this panel, reactions which are relevant only to PLCγ activation are represented inside the dashed boxes and are not included in the model for Sos activation.


A.
Box 1.
EGF
Y992
Y1068
Y1148/74
R
Ra
R2
1
2
3
4
Box 4.
PLCγ▪pY
PLCγP▪pY
PLCγ
PLCγP
5
6
7
Box 3.
Box 2.
ShcP
Grb2▪Sos
20
ShcP▪pY
Grb2
Sos
19
24
18
15
13
Shc
Sos
17
14
RP
RP1148/74
RP1068
RP992
11
9
10
B.
12 Grb2 + Sos ⟷ Grb2▪Sos
23 ShcP + Grb2▪Sos ⟷ ShcP▪Grb2▪Sos
22 ShcP▪Grb2 + Sos ⟷ ShcP▪Grb2▪Sos
21 ShcP + Grb2 ⟷ ShcP▪Grb2
16 ShcP ⟶ Shc
8 PLCγP ⟶ PLCγ
25 PLCγ ⟷ (PLCγ P)
C.
D.
pY
Yp
Yp▪Grb2
Sos▪Grb2▪ShcP▪pY
PLCγ▪pY
PLCγP▪pY
Grb2▪pY
Sos▪Grb2▪pY
pY▪Grb2
ShcP▪pY
pY▪Shc
pY▪ShcP
pY▪ShcP▪Grb2
pY▪ShcP▪Grb2▪Sos

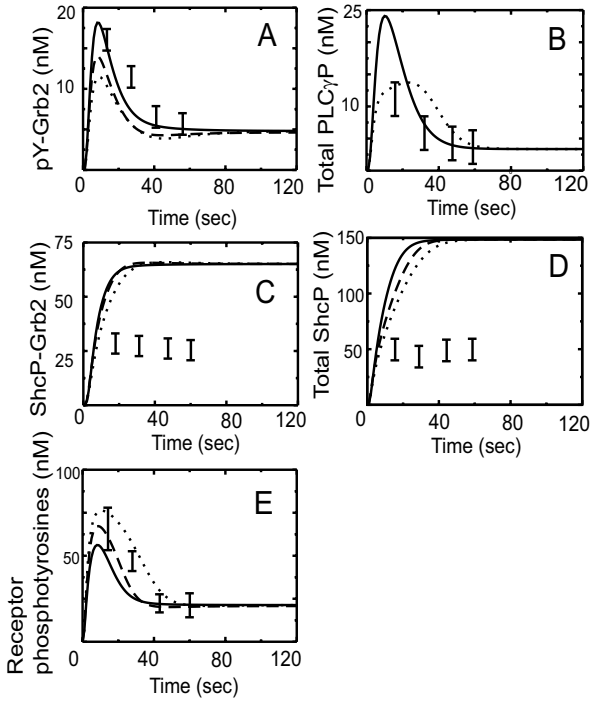


Figure 2. Predictions of the model of Kholodenko et al. (1999) (dotted line), the version of this model with PLCγ removed (broken line) and the network model (solid line). Experimental data of Kholodenko et al. (1999) are shown in each plot. (A) Grb2 bound to EGFR. (B) Total phosphorylated PLCγ. (C) Grb2 coprecipitated with Shc. (D) Total phosphorylated Shc. (E) Number of receptors with at least one phosphorylated tyrosine.

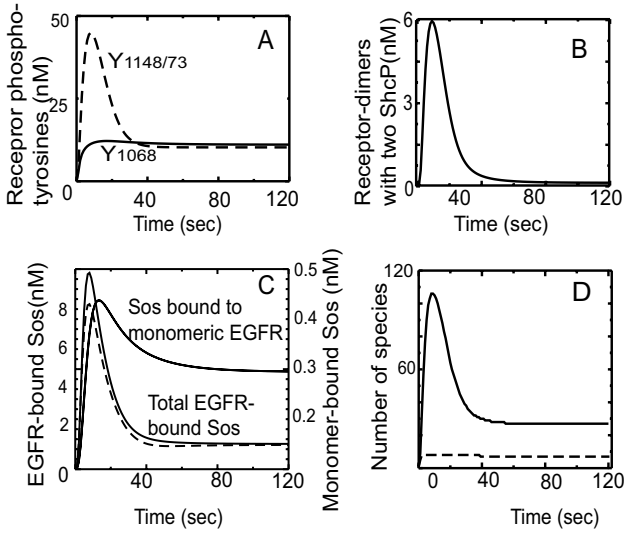


Figure 3. Some new predictions of the network model. (A) Different levels of phosphorylation for tyrosine residues that serve as binding sites for Grb2 (solid line) and Shc (broken line). (B) The amount of dimeric receptor species containing two phosphorylated Shc proteins. (C) Amount of Sos bound to monomeric receptors (right axis) and the total amount of Sos bound to receptors (left axis). The solid lines are derived from the network model; the broken line is derived from the version of the pathway-like model with PLCγ excluded. (D) Total number of receptor species containing 98% of the receptor mass as predicted by the network model (solid line) and the version of the pathway-like model with PLCγ excluded (broken line).

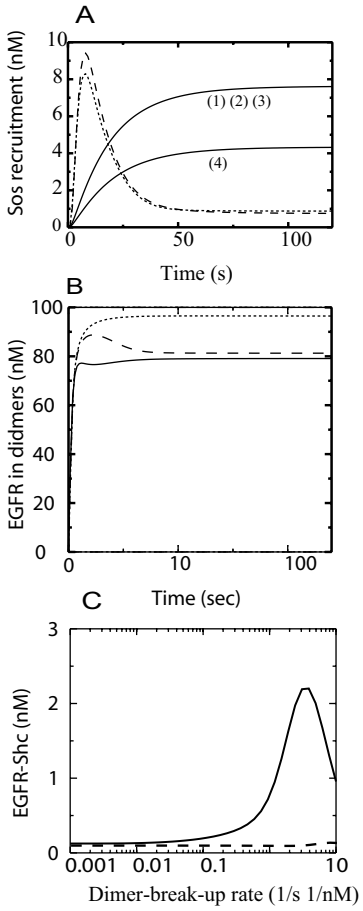


Figure 4. Comparing predictions of the two models for different hypothetical experiments. (A) Level of Sos recruitment for wild type (WT) and two mutations: knock out of EGFR tyrosines that bind Shc (EGFR-Y ko), and knock out of the Shc tyrosines (Shc-Y ko). For WT, the network model (broken line) and the pathway-like model (dotted line) express essentially the same level of Sos recruitment at all times. For the first mutation, the pathway-like (curve 1) and network (curve 2) models predict the same level of Sos recruitment. For the second mutation, predictions of the pathway-like and network models are different. The network model (curve 3) predicts the same level of Sos recruitment as for the first mutation, whereas the pathway-like model (curve 4) predicts a significantly lower level. (B) The number of receptor dimers depends on the total cytosolic Grb2 concentration in the pathway-like model (broken line represents the amount of receptor dimers for the nominal value of Grb2, dotted line represents the amount of receptor dimers for the 100-fold increase in cytosolic Grb2 concentration) but not in the network model (solid

line). (C) Shc recruitment as a function of the rate of dissociation of antibody-induced dimers of EGFR. Receptor dimerization is induced by a bivalent monoclonal antibody rather than by EGF. The prediction of the modified network model (solid line) and the modified pathway-like model (broken line) differ. These models are available as BioNetGen input files at our website (`http://cellsignaling.lanl.gov`).

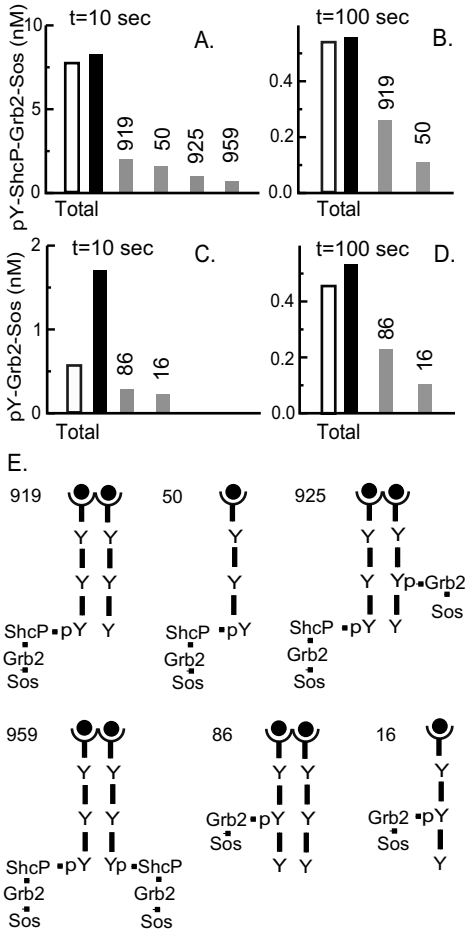


Figure 5. Distribution of Sos among receptor species. Black bars correspond to the network model; white bars correspond to the model of Kholodenko et al. (1999). Gray bars represent the receptor species of the network model containing more than 5% of Sos. These species are referred to by index and are graphically illustrated in panel E. (A) Distribution of Sos bound to pY1148/73 via ShcP-Grb2 at 10 sec. (B) Distribution of Sos bound to pY1148/73 via ShcP-Grb2 at 100 sec. (C) Distribution of Sos bound to pY1068 via Grb2 at 10 sec. (D) Distribution of Sos bound to pY1068 via Grb2 at 100 sec. (E) Species accumulating more than 5% of total Sos, which are referenced by index in the panels above.

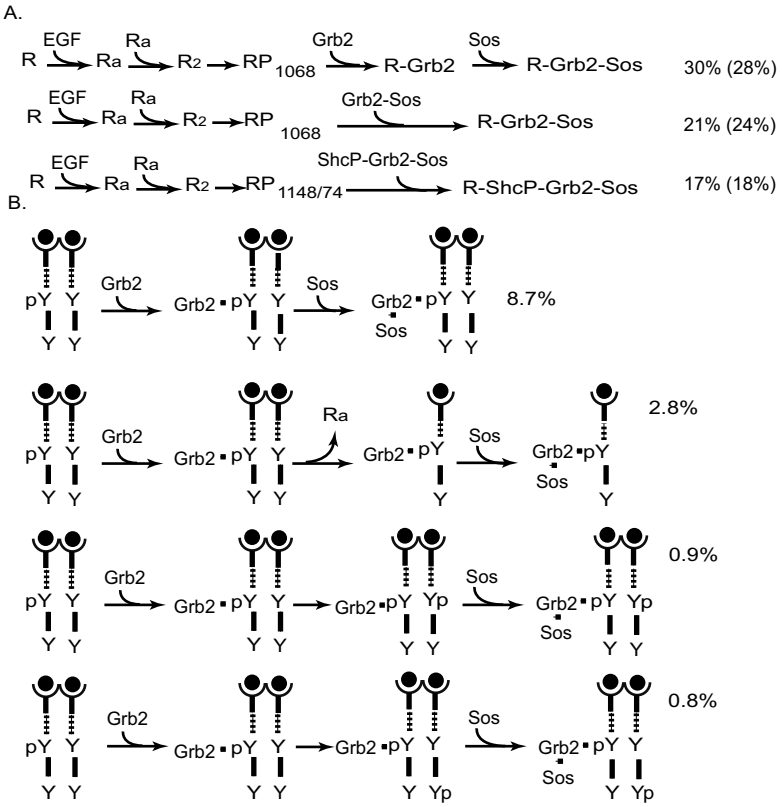


Figure 6. Prevalent reaction sequences leading to Sos recruitment at steady state. (A) Top three classes of reaction paths leading from a naked receptor to a receptor associated with Sos. The first two steps are EGF-EGFR binding and receptor dimerization. The third step is phosphorylation of a receptor tyrosine residue; different receptor phosphoforms are generated, illustrated in Box 2 of Figure 1A. After these steps, paths diverge. Percentages indicate the frequencies of the paths. Results for both models are given, with results derived from the pathway-like model in parentheses. (B) Top chemically distinct paths within the most prominent class of pathways leading to Sos recruitment, starting from a dimer in which pY1068 of one receptor is phosphorylated. The first three steps (EGF-EGFR binding, receptor dimerization, and tyrosine phosphorylation) are not shown. Although all the paths illustrated in panel B lead to formation of Grb2-Sos complex at pY1068, the receptor species involved differ significantly - they include both dimers and monomers and different phosphoforms of receptors. Percentages indicate the frequencies of these

specific paths among all paths leading to Sos recruitment, as derived from the network model. Results do not take into account modifications of the receptor tyrosine residue Y992, which is not involved in Sos activation.

Tables

Table 1. Reaction classes and parameter estimates associated with Sos recruitment. Signaling steps are illustrated in Figure 1. Initial concentrations are $[\text{EGFR}]_{\text{total}}$= 100 nM, $[\text{EGF}]_{\text{total}}$= 680 nM, $[\text{Grb2}]_{\text{total}}$= 58 nM, $[\text{Grb2-Sos}]_{\text{total}}$= 27 nM, $[\text{Shc}]_{\text{total}}$= 150 nM, $[\text{Sos}]_{\text{total}}$= 7 nM. Parameter values are taken from Kholodenko et al. (1999). Some of these parameters are based on experimental measurements, but most are estimates obtained through a fitting procedure. Binding and phosphorylation reaction rate constants are denoted as k_n and dissociation and dephosphorylation reaction rate constants are denoted as k_{-n}. First- and second-order rate constants are expressed in units of s^{-1} and $\text{nM}^{-1} \cdot \text{s}^{-1}$, respectively. Modifications of particular rate constants made in this study are marked with footnotes.

Step	Reaction class	Parameter values	Number of reactions
1	Ligand-receptor binding	$k_1 = 0.003; k_{-1} = 0.06$	48
2	Receptor dimerization	$k_2 = 0.01; k_{-2} = 0.1$	600
3	Receptor tyrosine phosphorylation	$^a k_3 = 0.5;\ ^b k_{-3} = 4.505$	96 + 144[c]
4	Receptor tyrosine dephosphorylation	$^b k_{-3} = 4.505$	104+156[c]
9	Binding of Grb2 to pY_{1068}	$^a k_9 = 0.0015; k_{-9} = 0.05$	24+288[d]
10	Binding of Sos to pY_{1068}-Grb2	$k_{10} = 0.01; k_{-10} = 0.06$	24+288[d]
11	Binding of Sos-Grb2 to pY_{1068}	$^a k_{11} = 0.00225; k_{-11} = 0.03$	24+288[d]
12	Cytosolic binding of Grb2 and Sos	$k_{12} = 0.0001; k_{-12} = 0.0015$	2
13	Binding of Shc to $pY_{1148-1173}$	$^a k_{13} = 0.045; k_{-13} = 0.6$	16+192[d]
14	Phosphorylation of Shc	$^a k_{14} = 3;\ ^a k_{-14} = 0.03$	8+192[d]
15	Binding of ShcP to $pY_{1148-1173}$	$^a k_{15} = 0.00045; k_{-15} = 0.3$	16+192[d]
16	Cytosolic ShcP dephosphorylation	$^b k_{-16} = 0.005$	1
17	Binding of Grb2 to $pY_{1148-1173}$-ShcP	$k_{17} = 0.003; k_{-17} = 0.1$	16+192[d]
18	Binding of ShcP-Grb2 to $pY_{1148-1173}$	$^a k_{18} = 0.00045; k_{-18} = 0.3$	16+192[d]
19	Binding of Sos to $pY_{1148-1173}$-ShcP-Grb2	$k_{19} = 0.01; k_{-19} = 0.0214$	16+192[d]
20	ShcP-Grb2-Sos binding to $pY_{1148-1173}$	$^a k_{20} = 0.00012; k_{-20} = 0.12$	16+192[d]
21	Cytosolic binding Grb2 to ShcP	$k_{21} = 0.003; k_{-21} = 0.1$	2
22	Cytosolic binding Sos to ShcP-Grb2	$k_{22} = 0.03; k_{-22} = 0.064$	2
23	Cytosolic binding Grb2-Sos to ShcP	$k_{23} = 0.021; k_{-23} = 0.1$	2
24	Binding of Grb2-Sos to $pY_{1148-1173}$-ShcP	$k_{24} = 0.009; k_{-24} = 0.0429$	16+192[d]
Total	20 reaction classes	37 parameters	3,749

[a] This rate constant is scaled by a factor of 0.5 to account for the different stoichiometry of binding sites in the model of Kholodenko et. al. (1999) and the network model (one binding site per dimer in the pathway-like model versus one binding site per receptor in the network model).

[b] This rate constant is that for a first-order reaction with a rate law of the form kC. This rate law replaces the Michaelis-Menten rate law in the original model of Kholodenko et al. (1999). This simplification, which does not affect results, is discussed in Section 2.2.

[c] The number of reactions involving Y1068 plus the number of reactions involving Y1148/73, respectively

[d] The number of reactions involving receptor monomers plus the number of reactions involving receptor dimers, respectively

Table 2. Reaction classes and parameter estimates associated with PLCγ activation. Reaction steps are illustrated in Figure 1. Initial concentrations are $[EGFR]_{total}$= 100 nM, $[EGF]_{total}$= 680 nM, $[PLC\gamma]_{total}$= 105 nM. Parameter values are taken from Kholodenko et al. (1999). Nomenclature of rate constants and units are the same as in Table 1.

Step	Reaction class	Parameter values	Number of reactions
1	Ligand-receptor binding	$k_1 = 0.003; k_{-1} = 0.06$	8
2	Receptor dimerization	$k_2 = 0.01; k_{-2} = 0.01$	20
3	Receptor tyrosine phosphorylation	$^a k_3 = 0.5;\ ^b k_{-3} = 4.505$	4
4	Receptor tyrosine dephosphorylation	$^b k_{-3} = 4.505$	6
5	Binding of PLCγ to Y_{992}	$^a k_5 = 0.03; k_{-5} = 0.2$	12
6	Transphosphorylation of PLCγ	$^a k_6 = 0.5;^a k_{-6} = 0.025$	10
7	Binding of PLCγP to Y_{992}	$^a k_7 = 0.003; k_{-7} = 0.3$	12
8	Dephosphorylation of PLCγ	$^b k_{-8} = 0.01$	1
25	Inactivation of PLCγP	$k_{25} = 1; k_{-25} = 0.03$	2
Total	9 reaction classes	15 parameters	75 reactions

[a] Same as footnote [a] in Table 1.

[b] Same as footnote [b] in Table 1